DOCUMENT RESUME

ED 378 226 TM 022 595

AUTHOR Borko, Hilda; And Others
TITLE Teachers' Ideas and Practices about Assessment and Instruction. A Case Study of the Effects of Alternative Assessment in Instruction, Student Learning and Accountability Practices.
INSTITUTION National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO CSE-TR-366
PUB DATE Oct 93
CONTRACT R117G10027
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Accountability; Attitude Change; Beliefs; Case Studies; *Educational Assessment; Educational Practices; *Elementary School Teachers; Grade 3; *Instruction; Interviews; Knowledge Level; Mathematics; Primary Education; Reading; *Teacher Attitudes; *Test Construction; Workshops
IDENTIFIERS *Alternative Assessment; Performance Based Evaluation

ABSTRACT

This report is one of a series that documents progress in understanding the effects of new assessment based on case studies in three elementary schools. It presents the preliminary analyses of interviews with teachers about their knowledge, beliefs, and practices related to assessment and instruction in reading and mathematics. Changes in beliefs and practices are also examined during a year-long intervention designed to help teachers develop performance assessments. Subjects were 3rd-grade teachers at three elementary schools who participated in weekly workshops in assessment development. It was clear from these preliminary findings that teachers added to their repertoires of assessment activities. In reading, they clearly considered these new activities--running records and written summaries--to be new forms of assessment, and they reported new insights about reading and student skills. In mathematics teachers did incorporate new activities, but only some of the teachers considered these activities to be assessment tools. Interview responses do not make clear the extent to which teachers are receiving new insight about student achievement from the new activities. End-of-year interviews may give more information about changes in teacher attitudes and practices. (Contains 8 references.) (SLD)

ED 378 226

CRESST

National Center for Research on Evaluation, Standards, and Student Testing

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Teachers' Ideas and Practices about Assessment and Instruction

A Case Study of the Effects of Alternative Assessment in Instruction, Student Learning and Accountability Practices

CSE Technical Report 366

Hilda Borko, Maurene Flory, and Kate Cumbo
CRESST/University of Colorado at Boulder

▸ UCLA Center for the Study of Evaluation

in collaboration with:

▸ University of Colorado
▸ NORC, University of Chicago
▸ LRDC, University of Pittsburgh
▸ The RAND Corporation

TM022595

BEST COPY AVAILABLE

Teachers' Ideas and Practices about Assessment and Instruction

A Case Study of the Effects of Alternative Assessment in Instruction, Student Learning and Accountability Practices

CSE Technical Report 366

Hilda Borko, Maurene Flory, and Kate Cumbo
CRESST/University of Colorado at Boulder

October 1993

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Graduate School of Education
University of California, Los Angeles
Los Angeles, CA 90024-1522
(310) 206-1532

Copyright © 1993 The Regents of the University of California

This research was supported in part under the Educational Research and Development Center Program cooperative agreement R117G10027 and CFDA catalog number 84.117G as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

PREFACE

The current intense interest in alternative forms of assessment is based on a number of assumptions that are as yet untested. In particular, the claim that authentic assessments will improve instruction and student learning is supported only by negative evidence from research on the effects of traditional multiple-choice tests. Because it has been shown that student learning is reduced by teaching to tests of low level skills, it is theorized that teaching to more curricularly defensible tests will improve student learning (Frederiksen & Collins, 1989; Resnick & Resnick, 1992). In our current research for the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) we are examining the actual effects of introducing new forms of assessment at the classroom level.

Derived from theoretical arguments about the anticipated effects of authentic assessments and from the framework of past empirical studies that examined the effects of standardized tests (Shepard, 1991), our study examines a number of interrelated research questions:

1. What logistical constraints must be respected in developing alternative assessments for classroom purposes? What are the features of assessments that can feasibly be integrated with instruction?

2. What changes occur in teachers' knowledge and beliefs about assessment as a result of the project? What changes occur in classroom assessment practices? Are these changes different in writing, reading, and mathematics, or by type of school?

3. What changes occur in teachers' knowledge and beliefs about instruction as a result of the project? What changes occur in instructional practices? Are these changes different in writing, reading, and mathematics, or by type of school?

4. What is the effect of new assessments on student learning? What picture of student learning is suggested by improvements as measured by the new assessments? Are gains in student achievement corroborated by external measures?

5. What is the impact of new assessments on parents' understandings of the curriculum and their children's progress? Are new forms of assessment credible to parents and other "accountability audiences" such as school boards and accountability committees?

This is one of four reports that document our progress in understanding these questions, based on case studies in three elementary schools.

TEACHERS' IDEAS AND PRACTICES ABOUT ASSESSMENT AND INSTRUCTION[1, 2]

Hilda Borko, Maurene Flory, and Kate Cumbo
CRESST/University of Colorado at Boulder

This paper presents results of preliminary analyses of interviews with teachers about their knowledge, beliefs, and practices related to assessment and instruction in reading and mathematics. In particular, we describe and examine teachers' knowledge and beliefs about what it means to know reading and mathematics, and their ideas and practices related to assessment and grading in the two subject areas. We also examine the nature and extent of changes in these characteristics during the first semester of a year-long intervention designed to help the teachers develop performance assessments in reading and mathematics.

The literature is replete with examples of how difficult it is for teachers to change their knowledge, beliefs and practices in fundamental ways (e.g., McDiarmid, 1990; Putnam, Heaton, Prawat, & Remillard, 1992; for a review, see Borko & Putnam, in press). Yet, many current educational reform movements are calling for just such fundamental changes. The performance assessment movement is no exception. Teachers are being asked to reject traditional forms of classroom assessment, such as the objective, paper-and-pencil unit tests and chapter tests typically found in mathematics and reading textbooks, and to replace them with performance assessments. The claim is that these more "authentic" performance assessments, with their focus on problem solving and higher order thinking skills, will improve teaching and enhance student learning (Resnick & Resnick, 1992; Wiggins, 1989).

---

[1] This paper was presented at the annual meeting of the American Educational Research Association, Atlanta, GA, April, 1993.

[2] This research is part of a larger research project, Studies in Improving Classroom Teaching and Local Assessments. All members of the research team contributed to the conceptualization and data collection for this paper. We extend special thanks to the teachers who are participating in the project. Without their support, commitment, and countless hours spent in workshops and interviews with us, this project would not have been possible.

For teachers to incorporate such fundamental changes into their ongoing instructional practices, they must have the knowledge necessary to implement the changes and the beliefs to support them. Given the literature on teacher change in other areas, it seems reasonable to expect that such transformations in teachers' ideas and practices about assessment might be difficult to accomplish. This expectation receives some support from members of the assessment community such as Stiggins (1991), who warns that many teachers lack the "assessment literacy" necessary to implement performance assessments.

Our research team's commitment to classroom-based performance assessments, coupled with our concerns about teachers' ability to implement such assessments in their classrooms without support and guidance, were prime motivators for the larger project of which this study is a part. The specific commitment of this paper's authors to understanding and improving the process of learning to teach led to this particular study of changes in the ideas and practices of the teachers participating in the project.

Methods and Data Sources

This study is one component of a 3-year research project. The major purposes of the larger project are to work with third-grade teachers to develop classroom performance assessments in reading and mathematics, and to study the effects of performance assessments on teachers' knowledge and beliefs, instructional practices, and student learning. The project uses a "bottom-up" intervention, in which the research team is meeting with teachers on a weekly basis to provide support and guidance in the development of classroom performance assessments. Our intention in these weekly workshops is to foster changes in the teachers' assessment practices by helping them to think about instructional goals, develop or select assessment tasks appropriate to those goals, and articulate scoring criteria.

Participants

Primary participants in the project are the third-grade teachers at three elementary schools in a school district on the outskirts of Denver. Schools wishing to be considered for the project submitted a short proposal and documentation of approval by the principal, all third-grade teachers, and the

parent accountability committee. Participating schools were selected to represent a range of socioeconomic levels and student achievement.

Data Sources

Data sources for this paper are interviews with the third-grade teachers conducted before we began our weekly workshops (September 1992) and again at the end of the fall semester (January 1993). Semi-structured questions focused on teachers' knowledge, beliefs and practices related to assessment, instruction, and the relationship between assessment and instruction. Interviews were conducted by members of the research team, at the participating schools, during the school day. All interviews were audiotaped and transcribed.

Data Analysis

Our analyses for this paper examined most closely teachers' responses to a subset of interview questions that addressed their conceptions of what it means to know mathematics and reading, their reports of assessment and grading practices, and their beliefs about the advantages and limitations of performance assessments. We also considered participants' responses to other interview questions, when they were relevant to the issues that were the focus of this paper.

We developed a coding system based on our research questions and an initial reading of a sample of fall interview transcripts. The system included categories such as "know-math," "how assess-math," "why assess-math," "grade-math," and their parallels for the subject area of reading. We added several categories that focused on change (e.g., "change know-math," "change assess-math") after reading a sample of the winter interviews.

Our analysis consisted of coding the data separately for each school, identifying major themes in the data, writing cases of the individual schools organized around these themes, and then conducting a cross-case analysis to identify similarities and differences in the patterns of knowledge, beliefs, and practices that characterized teachers at the three schools. This paper presents the cross-case analysis. We conclude with a discussion of the nature and extent of change in the teachers' ideas and practices, and with some hypotheses about factors that may explain patterns in our findings.

Patterns in Teachers' Knowledge, Beliefs, and Practices

Our cross-case analysis begins with teachers' knowledge, beliefs, and practices related to reading, and then examines parallel issues for the subject area of mathematics. For the most part, we focus on school-level patterns and trends. Whenever individual teachers provide particularly important insights into the change process, we discuss them as well.

Reading Programs

At a very general level, teachers in the three schools held a shared conception of what it means to know reading. The two major dimensions that comprised their conceptions were skill and understanding. Affect played a consistent, but much less prominent, part in their responses.

Teachers at all three schools identified fluency and word-level reading strategies as important components of the skill dimension at the beginning of the school year. Abby's (all names used in reference to the project are pseudonyms) description of one characteristic of excellent readers captures their views: Excellent readers "would know what to do when they came to a word that they didn't know. They'd have the skill to figure that out; they'd use all sorts of strategies . . . that we, as readers, use and don't even think about."

The ability to identify story elements was seen as a central component of "understanding" by teachers at all three schools. At a more fine-grained level of analysis, however, conceptions of this component of reading differed across the schools. Teachers at Walnut School wanted students to "identify story elements—character, setting, events, problems, inference based on actions or personalities of characters." Teachers at Spruce School talked in vaguer terms about their desire that students be able to "make sense" out of a story by identifying story elements. At Pine School, teachers talked about identification of story elements as if it were a skill to be mastered in a fairly routinized manner. To facilitate mastery, they had developed a sequence of structured tasks to help students move from responding to a series of questions asking for specific words or phrases in the text that identified each story element, to writing a paragraph in which story elements were reported.

Other components of reading for understanding that were mentioned by teachers differed across schools. At Walnut, the ability to tell others about what

one has read was a prominent element of understanding. Teachers at Pine talked about reading for meaning beyond the literal aspects of the text. Teachers at Spruce seemed to consider word-level meanings, in addition to "making sense" out of a story, as a component of understanding.

At a general level, fall interviews also revealed similarities in the assessment and grading practices at the three schools. Teachers at all three schools assessed for both skills and understanding, and they all used paper-and-pencil tasks as well as observation and listening approaches.

Again, differences emerged when a more fine-grained analysis was conducted. These differences were associated with one basic difference in the schools' reading programs. Reading at Walnut and Spruce was organized around a Reader's Workshop approach and, for the most part, children read in self-selected trade books. In accord with Reader's Workshop, a major portion of assessment consisted of individual conferences with students about the books they were reading. All teachers at the two schools had developed systems for keeping written records of the conferences.

For teachers at Spruce, assessment consisted primarily of "listening," "observing," and "looking at student writing." The teachers seemed to place considerable emphasis on fluency and word-level understanding as they listened to students and asked questions about their reading. All of the teachers kept written records of conferences, although there were individual differences in the extensiveness of these records.

Teachers at Walnut were much more explicit in describing both what they looked for in their conferences and what information they recorded. As Abby explained, referring to how she determined what her students know and can do in reading, "Most of it is by listening to them read, and talking to them, conferencing with them, asking direct questions . . . that give you a feel for what the kids are keying into, or if they can key into character development or anything like that." For Jackie, conferences also provided information about the affective component of reading: "Are they enjoying the book? . . . Do they like it?" Although their specific record-keeping systems differed, the teachers seemed to be keeping track of similar kinds of information. For example, Jackie kept anecdotal records in a notebook, with information such as the book a child is reading, his or her reaction to that book, oral reading proficiency, and areas in

need of instruction. "In the past, I have had a notebook that has tabs, one for each child. So that I have one notebook, and everybody's name is there, and there's notebook paper in there, and basically it's just keeping an anecdotal record . . . kind of tracking what the kids are doing." She tried to keep these records "while I'm conferencing. . . . that's the best time to do it, because it's fresh in your mind, you have 26 students you're working with, and you just can't remember which student needs work in oral expression, and which student needs help in self-selection, and where the kids are."

In contrast to both schools, Pine teachers used a basal textbook series and placed students in ability groups. For assessment purposes, these teachers placed more emphasis on traditional paper-and-pencil tasks and written work than did their colleagues at Walnut and Spruce. Although they acknowledged the importance of anecdotal records, only half of them had devised systems for keeping track of these records, and none seemed comfortable with any strategy they knew of for taking the records into account when determining grades. The overriding characteristic of their view of reading assessment was its problematic nature. As Elly noted, "I think reading is extremely hard to evaluate."

Despite these differences across the schools, a key theme among all teachers was their desire to learn more, through work on the CU project, about keeping informal records during the reading program.

Changes in reading programs. Teachers at all three schools had incorporated two types of activities into their reading programs—running records and written summaries—by the time of the winter interviews. Both of these activities were introduced by the CU project's literacy expert, at workshops early in the year, as potential assessment tools. Perhaps the most consistent change associated with participation in the CU project, in either reading or mathematics, was that all teachers were using running records (although to varying degrees) with their below-grade-level readers. Teachers talked about their use of running records in very similar ways. Most teachers at each school reported that the use of running records increased their awareness of students' reading abilities. Many of the teachers reported that information they acquired from running records informed their instruction. Sara's comment captured both of these ideas: ". . . things like running records . . . tend to tell you what types of errors they are making, and they do change my instruction as to what I need to focus on with these kids." Perhaps one reason running records were adopted by

the teachers is that, as one teacher explained, they are "a quick way to follow the kids' progress," and they "are much easier than any other kind of reading test I've found."

All of the teachers also reported that they were using written summaries as part of their reading assessments. As was the case with running records, they talked about written summaries, and apparently used them, in similar ways. All teachers developed scoring criteria for summaries (typically, in conjunction with their teammates, and in part during their biweekly workshops with CU project team members). They shared these criteria with their students and, to varying degrees, students participated in the evaluation of their own written summaries. Teachers at all three schools reported an increased awareness of their students' understanding of text. They also reported that students appeared to be more aware of their own reading ability and their teachers' evaluation criteria, and better able to evaluate their reading performance (i.e., their written summaries). As one teacher at Walnut explained, the students gained "an understanding of what teachers expected from them. . . . I don't think some kids ever had a clue as to how we got their grades in reading."

We did find several differences across schools in their discussions of running records and written summaries. In general, teachers at Walnut and Pine were more explicit than their colleagues at Spruce in their discussions of the specific kinds of information revealed by these new assessment tools. As one teacher at Walnut explained, running records "tell us specifically, maybe some sounds and letter blends that kids don't know. . . . They can also tell us which kids recognize their errors and do something about them."

Teachers at Walnut discussed student changes to a greater extent than did teachers at the other two schools. They addressed affective (e.g., feeling better about themselves as readers) as well as cognitive (e.g., use of known spelling patterns, understanding of teachers' expectations and grading practices) dimensions of student change. For example, Janet noted about the below-grade-level readers in her class and their experiences with running records, "I think kids that are struggling can look at those [words that they read correctly] and say, 'look at all these words I know,' and kind of get away from 'I am a bad reader'."

Also, as a result of their experiences with running records and new insights about their students' reading abilities, teachers at Walnut decided to study the Benchmark reading program as a team, in order to learn specific word-level strategies to incorporate into their instructional programs. Judy talked about the differences that these experiences had made for her. "In the fall . . . I did not feel that I knew how to meet the needs of the first-grade readers. And I was running out of material for them to read. . . . And I didn't have the children that I had pushed beyond third-grade level, because I wasn't aware of what they were able to do." By the time of the winter interview, she was "helping some of those kids who are having reading difficulty . . . by looking at the words that they are having difficulty with and picking out the small key spelling patterns and sounding them out that way."

Another difference between schools was that teachers at Walnut viewed using written summaries as a way of assessing reading competence to be problematic, to some extent. Their concern was that some students are proficient readers, but are not able to write good summaries. They were unwilling to penalize students' reading grades because of weaknesses in writing. As Janet explained, "The summaries, we took into consideration, but certainly did not use those as a major part of the reading grade because we have good readers who read well, who recognize nonsense, who understand stories and reading, who were not writing great summaries. And we did not feel that we could grade them low." The teachers at Spruce, in contrast, saw no problem in assessing reading proficiency with techniques that incorporate writing skills. All Spruce teachers seemed to view reading and writing as going "hand-in-hand." Two talked explicitly about the connection between reading and writing. Teachers at Pine did not address this issue.

Finally, one aspect of individual differences among teachers is important to note. For at least a small number of teachers, changes in their assessment and instructional practices seemed to be accompanied by changes in their conceptions of reading. Janet's comment represents such a shift: "We have made an error in assuming that all kids learn to read from immersion in literature and building vocabulary. And that works for a lot of kids, but not for all." This insight led Janet to depart from some of the tenets of a Reader's Workshop approach to teaching reading, and to "incorporate a little more phonics" and explicit instruction in strategies such as compare/contrast and key spelling

patterns in a school where "it seems like that for a few years, 'phonics' was a four-letter word." In contrast, other teachers saw similar changes in assessment and instructional practices as the addition of new strategies into their repertoires, but considered these additions to be consistent with their existing conceptions or philosophies.

Mathematics Programs

As was true with reading, at a general level, teachers' conceptions of what it means to know mathematics were similar across the three schools and included skills, understanding, and (to a lesser extent) affective components. Differences in emphasis across schools were greater in mathematics than in reading. For teachers at Pine and Spruce, the primary emphasis in instruction and assessment, at least for the first quarter, was on math facts and skills. All of the teachers at Pine stressed that "good mathematicians know their facts." A common goal for the first grading period is "making sure that they are memorizing their math facts through 18."

Teachers at Walnut acknowledged that students must "master the facts," and that computation and number facts are components of "knowing mathematics," but they found the focus on skills to be problematic. Janet was among those most hesitant in attributing importance to mathematics skills. Although she acknowledged that "we do spend a little bit of time in the room working on mastering the facts," and she was not ready to say "that algorithms . . . shouldn't be taught," she was quick to point out that "the computation that we do is really a means to an end." That end, for her, was mathematical understanding and problem solving.

At all three schools, the meaning of "understanding mathematics" was ambiguous. In its fullest sense, understanding seems to include the ability to explain steps in a mathematical procedure (for example, the ability "to explain and describe how they derive an answer") as well as understanding the conceptual underpinnings of mathematical procedures and the connections among concepts and procedures (for example "number sense," "recognizing patterns," or "seeing connections"). It typically was not possible to determine whether and to what extent individual teachers' conceptions included both components. This was less true at Walnut, where teachers' comments were more extensive, and where their conceptions appeared to incorporate both

components. For example, when asked what it means to be excellent in mathematics, Abby responded ". . . that they have a good understanding of numbers and when you would use addition, when you would use subtraction. . . . They can see a connection between patterns and a connection between numbers. . . . They can explain how they came up with an answer, maybe even give more than one way to solve . . . "

Teachers' conceptions of problem solving were also ambiguous. To the extent that we can tell, however, they seemed to differ across the schools. Teachers at Walnut and Spruce did little more than mention problem solving as one aspect of what it means to know or be excellent in mathematics. For example, students who are excellent in mathematics can "use information from many sources to apply to the problem," and "know what resources to go get, to help [solve a problem]." Teachers at Pine were somewhat more explicit in their comments. Interestingly, more than their colleagues at the other schools, Pine teachers seemed to focus on problem solving as a set of procedures and steps that can be taught and learned in a fairly routinized manner. This orientation to problem solving was similar to their view of summarizing in the subject area of reading.

In mathematics, in contrast to reading, assessment strategies at all three schools focused on skills to a greater extent than understanding. All of the teachers were fairly dependent on traditional forms of assessment, particularly when it came to assigning grades in mathematics. And, none of the teachers had well-developed systems for keeping track of information gathered while observing and listening to students during math lessons.

Beyond these general patterns, however, school-level differences in assessment strategies were greater in math than in reading. Timed tests in the four basic operations are mandated by the school district and administered at all three schools. They were much more prominent at Spruce than at Pine or Walnut, both in the teachers' interview responses and in their mathematics programs. In fact, teachers at Spruce generally were more focused on skills and facts than were their colleagues at the other schools. They were also much vaguer in their descriptions of other forms of assessment they use.

Teachers at Pine relied almost exclusively on other paper-and-pencil tasks to supplement timed fact tests for assessment and grading purposes. They were

much more explicit than their colleagues at Walnut or Spruce, in describing the various types of tasks they use; their repertoire included daily written work (usually problems copied from the book), chapter tests from the books, and personally-created end-of-unit tests.

In contrast, observation and listening seemed to play a much more prominent role in the assessment systems of teachers at Walnut (though not in their grade determination). More than teachers at the other schools, they talked explicitly about what they look and listen for when observing students engaged in mathematical activities. Jackie was the most specific. She did "a lot of observing what they need to complete a task and how successful they are with that task. . . . those students who are frustrated, who don't understand and who perhaps are relying on a partner or relying on the teacher for procedures." One specific focus of her observations was students' use of manipulatives. She would "give them a task and give them the option of using manipulatives or not; it becomes quite apparent those students who do not need the manipulatives." She gave several examples of assessment for specific topics; for place value, addition, and subtraction: "I am watching them and observing how quickly they're able to set up a problem or work a problem with the beans, and then again describe what they're doing and how they're approaching it."

Changes in mathematics programs. In contrast to the subject area of reading, there were more school-level differences than similarities in the nature and extent of changes in mathematics assessment. There were, however, some consistencies across the schools. For example, the mathematics educator on the research team supplied the teachers with many suggestions for mathematics activities that use manipulatives and focus on higher order thinking and problem solving. Teachers at all three schools incorporated these activities into their mathematics programs.

By the midyear interviews, teachers at the three schools were asking students to provide written explanations for some of the math problems they solved. Similarly to the written summaries they incorporated into the reading assessments, they developed scoring criteria for these explanations, taught the criteria to students, and had students participate in the scoring process. They also reported that they knew more about their students' understanding of mathematics than was the case when they relied exclusively on more traditional forms of assessment. Jackie's comment, though more explicit than most,

represents the teachers' sentiment: "I think it [new types of assessment] gives us insight into the student's thinking and learning, more so than just a score or plus number. So I think what we have done with the project has really taught us as educators to question kids to think about how did you get that solution or what process did you follow. . . . I certainly learned . . . I think I know more about my students and where they're coming from than just are they right or wrong."

The schools differed, however, in the extent to which they used these new activities and explanations for assessment purposes. Teachers at Walnut clearly considered the activities and explanations to be part of their assessment systems. Each teacher developed a record-keeping system for the new activities, which included information about both written assignments and teacher observations. Janet used checks, plusses, and minuses. As she explained, the activities "lent themselves more to a check system" than to percentage scores or grades. Jackie kept track during observations with "a check sheet with everybody's name on it, and at the top, just the skill I'm looking for. And I would note whether they were showing mastery or needed more work." She used a 3-point numerical system to score written problem-solving assignments. Most teachers reported relying less on percentage scores on traditional paper-and-pencil tests, and more on these records of student performance, when evaluating students' mathematics proficiency and assigning report card grades.

The situation was more complex at Pine and Spruce. Teachers at these schools clearly considered the activities to be part of their mathematics programs. They talked about them, however, primarily as instructional tools. The activities were seen as important means for engaging students in higher order thinking and getting them to "communicat[e] their answers in language." The teachers acknowledged that they were learning more about their students and their mathematical understanding by using these activities. For example, "I think it tells me much more about the whole student as a learner, as a thinker. [Before] the only thing I looked at was if they knew their facts." However, most teachers continued to assess students with the traditional chapter tests and timed tests. Our sense is that, at the time of the winter interviews, Pine and Spruce teachers viewed the new components of their math programs—with the exception of written explanations—as instructional activities which provided

them with new insights about their students, rather than as assessment strategies *per se*.

Finally, as was the case in reading, we found some evidence that for at least a small number of teachers, the changes in their instructional and assessment practices seemed to be accompanied by fundamental changes in their conceptions of mathematics. Judy presents the clearest example, as reflected in her comment that "I'm looking at math differently." In the fall, her response to the question about what it means to be excellent in mathematics mentioned, in general terms, aspects of understanding such as "manipulate numbers," "estimate," and "having a feeling for numbers . . . and what they meant." At midyear, her answer was much more detailed, and she was more explicit about mathematical processes such as "explain how they got an answer," and "analyze her own mathematical thinking." She also expressed the belief that children can be excellent in mathematics because they "do things quickly and easily and accurately" or because they "ask great questions and [are] very thoughtful about problem solving and what [they do]. . . . So there are really two kinds of mathematicians."

Our sense is that Judy is an exception to the rule, and that most teachers did not experience fundamental changes in their conceptions of mathematics. As Janet explained, "I don't think that for me there has been a big philosophy change from last year to this year. . . . The project didn't give me a new philosophy. What it has done is push me on a little more in the philosophy I already had."

Speculations About Change

At this point in the project, it is premature to draw conclusions about the nature of changes in the teachers' knowledge, beliefs or practices that are resulting from their participation in the project. At most, we can offer some observations and tentative speculations about the future.

It is clear that in both reading and mathematics, teachers have added to their repertoires of activities. In reading they clearly consider these activities—running records and written summaries—to be new forms of assessment. Further, they reported new insights about students' reading, and they seem to have greater understandings, at least about below-grade-level readers' skills,

based on these assessments. One thing that is less clear, and that appears to be less consistent across schools and teachers, is the extent to which teachers have changed their instructional programs to take advantage of these new assessments and insights.

In mathematics, teachers have consistently incorporated new activities into their instructional programs. However, only some of the teachers consider these activities to be assessment tools. Also, the extent to which the activities have provided the teachers with greater insights into their students' mathematical understandings is not clear from their interview responses.

What is perhaps most unclear to us at this point in the project is the extent to which teachers' conceptions of reading, learning to read, mathematics, and learning mathematics have changed as a result of their experiences on the project and the new tools they have incorporated into their programs. We have evidence that a few of the teachers have begun to reconsider their conceptions. Most of the teachers, however, seem to have incorporated new ideas from the project into their existing reading and math programs, without making fundamental changes in these programs or in the conceptions of reading and math that underlie them. These patterns are neither surprising nor discouraging to us, given what the literature suggests about difficulties in effecting teacher change. Further, given the degree of change we have seen in some schools and some individual teachers, we are optimistic that if the teachers continue to make changes in their programs, to reflect on the impact of these changes on student learning, and to share their experiences in the weekly workshops, we will see more extensive changes in their conceptions by the end-of-year interviews.

References

Borko, H., & Putnam, R. (in press). Expanding a teacher's knowledge base: A cognitive psychological perspective on professional development. In T. Guskey & M. Huberman (Eds.), *New paradigms and practices in professional development*. New York: Teachers College Press.

Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, *18*(9), 27-32.

McDiarmid, G.W. (1990). Challenging prospective teachers' beliefs during early field experience: A quixotic undertaking? *Journal of Teacher Education*, *41*(3), 12-20.

Putnam, R., Heaton, R.M., Prawat, R.S., & Remillard, J. (1992). Teaching mathematics for understanding: Discussing case studies of four fifth-grade teachers. *Elementary School Journal*, *93*, 213-228.

Resnick, L.B., & Resnick, D.P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B.R. Gifford & M.C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Boston: Kluwer Academic Publishers.

Shepard, L.A. (1991). Will national tests improve student learning? *Phi Delta Kappan*, *73*, 232-238.

Stiggins, R. (1991). Assessment literacy. *Phi Delta Kappan*, *72*, 534-539.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, *71*, 703-713.